

# Une ontologie linguistique au service de la didactique du lexique

Ophélie Tremblay

Université du Québec à Montréal  
&

Centre de recherche interuniversitaire sur la formation et la profession enseignante (CRIFPE)  
tremblay.ophelie@uqam.ca

Alain Polguère

Université de Lorraine & ATILF CNRS  
alain.polguere@univ-lorraine.fr

## 1 Introduction

Cet article propose de revisiter les fondements linguistiques de la didactique du lexique. Plus particulièrement, il vise à mettre en évidence en quoi la Lexicologie Explicative et Combinatoire (Mel'čuk et coll. 1995) est un support théorique et descriptif approprié pour guider la spécification des notions linguistiques de base de l'enseignement du lexique, à travers une perspective intégrée dans laquelle les connaissances sur le lexique sont mises en relation avec les connaissances sur la syntaxe et la morphologie de la langue. Pour appuyer notre propos, nous présentons une ontologie des savoirs lexicologiques issus de la Lexicologie Explicative et Combinatoire, l'ontologie GTN. Cette ontologie a servi de fondement à l'élaboration d'un module de cours en didactique du lexique français destiné à de futurs maîtres au primaire au Québec (Tremblay, 2009), de même qu'à la rédaction d'un manuel d'enseignement du vocabulaire (Tremblay & Ancil, en préparation).

L'article se divise en trois sections. Dans la première section, nous présentons le contexte didactique et linguistique dans lequel s'inscrit notre recherche. Nous détaillons dans la deuxième section les caractéristiques principales de notre ontologie des savoirs lexicologiques en mettant en lumière leur portée didactique, notamment à travers des exemples d'utilisation de l'ontologie pour l'élaboration d'un module de cours en didactique du lexique. Enfin, nous précisons les retombées d'une telle recherche pour la formation des élèves et des enseignants de français ainsi que pour la composition et l'organisation des programmes scolaires d'enseignement du français.

## 2 De l'enseignement du « vocabulaire » à l'enseignement des savoirs lexicologiques

L'enseignement/apprentissage du vocabulaire, et à travers lui le développement de la compétence lexicale des élèves, est l'objet d'étude principal de la didactique du lexique. Mais la question des notions qui structurent la connaissance du lexique (polysémie, dérivation sémantique, locutions, collocations, etc.) constitue aussi un des axes de recherche de la didactique du lexique (et un des objets d'étude que devrait couvrir l'enseignement du vocabulaire, tant au primaire qu'au secondaire). C'est essentiellement dans cet axe que s'inscrit la réflexion que nous proposons dans cet article.

### 2.1 Importance de la compétence lexicale

Pour enseigner le français au primaire et encadrer le développement des compétences linguistiques de l'élève, l'enseignant doit non seulement faire preuve d'un excellent niveau de compétence linguistique, notamment d'un vocabulaire étendu, mais aussi acquérir les connaissances métalinguistiques nécessaires et suffisantes pour enseigner la langue en tant que système de règles :

The rules governing language must be a vital part of the teacher's knowledge ; otherwise pupil's engagement with language, whether it be in spelling or in poetry or in the lived experience of the individual, can never really be progressive, systemically developed, or properly evaluated and assessed (Carter, 1982 : 8).

Parmi les connaissances métalinguistiques à acquérir, les connaissances sur l'organisation et le fonctionnement du lexique jouent un rôle essentiel puisqu'elles permettent à l'enseignant d'encadrer chez ses élèves le développement d'une compétence fondamentale : la compétence lexicale. La compétence lexicale – la connaissance des unités lexicales d'une langue et la capacité à les utiliser et à les comprendre (les aspects productifs/actifs et réceptifs/passifs de la maîtrise du vocabulaire) – joue un rôle prépondérant dans les processus de communication à l'oral comme à l'écrit (Meara, 1996 ; Polguère & Tremblay, 2003 ; Simard, 1994). Plusieurs recherches ont d'ailleurs démontré que le vocabulaire constitue un facteur clé dans la réussite scolaire (Biemiller, 2011 ; Pagani et coll. 2011 ; Desrosiers & Ducharme, 2006), ce qui justifie qu'on porte une attention particulière à son enseignement et à son apprentissage, et ce, dès le primaire.

De plus, le lexique entretient des liens étroits avec la grammaire de langue, ce qui fait que les connaissances lexicales (les composantes centrales de la compétence lexicale) peuvent tout à fait être enseignées à travers une mise en relation avec les connaissances grammaticales. Un tel type d'enseignement est d'ailleurs préconisé par Jacqueline Picoche (2011), qui propose des principes fondateurs pour développer le vocabulaire en classe, notamment : 1) de donner la priorité au verbe, du fait notamment qu'il structure la phrase et permet d'explorer des combinaisons de mots et de travailler sur les dérivés et les familles de mots et 2) de ne pas séparer vocabulaire et syntaxe. Pour l'instant, il semble que l'application de ces principes ne soit pas pratique courante en classe. En effet, bien que la compétence lexicale joue un rôle de premier plan dans la communication langagière et que, de façon parallèle, les connaissances lexicales soient considérées comme l'une des composantes les plus importantes de l'acquisition du langage (David, 2000), les cours de langue en français langue première au primaire (en France aussi bien qu'au Québec) privilégient surtout l'étude de la grammaire ou encore de l'orthographe (David, Paveau & Petit, 2000). Cela fait en sorte que le développement de la compétence lexicale des élèves est négligé.

## 2.2 Place du lexique dans l'enseignement de la langue première au primaire

Certains facteurs peuvent expliquer le rôle périphérique du lexique dans l'enseignement de la langue première au primaire. D'abord, la compétence lexicale repose sur des connaissances lexicales organisées en un système extrêmement riche et complexe. En effet, les unités lexicales se caractérisent par trois propriétés fondamentales : un sens, une forme et des propriétés de combinatoire. La forme est la propriété des unités lexicales la plus travaillée à l'école, dans le cadre de l'apprentissage de l'orthographe lexicale. Le sens fait également l'objet d'un certain enseignement, mais celui-ci a lieu de façon plus occasionnelle que systématique, lors de la lecture, principalement. Quant aux propriétés de combinatoires (les informations qui permettent de savoir comment utiliser une unité lexicale en situation de discours : genre des noms, forme plurielle irrégulière des noms et des adjectifs, régime syntaxique, collocations, etc.), elles font rarement l'objet d'un enseignement explicite, malgré leur richesse et le fait qu'elles représentent un aspect essentiel de la « signature individuelle » de chaque unité lexicale.

Ensuite, les études s'étant intéressées à la façon dont les enseignants travaillent le lexique en classe (par exemple Anctil, 2011 ; Dreyfus, 2004) rendent compte du fait que ceux-ci le font de façon isolée, en s'attachant seulement à un petit nombre de savoirs, dans certains cas parce qu'ils ne disposent pas de connaissances métalinguistiques suffisantes. La représentation selon laquelle le lexique ne s'enseigne pas parce qu'il est trop vaste et peu organisé (Leeman, 2000) semble également encore tenace chez les enseignants, même si, comme le dit Picoche, « il ne faut pas se laisser décourager par l'immensité du lexique. En effet, les mots n'ont rien d'une masse informe » (p.1).

Enfin, les programmes de formation qui servent de référence aux enseignants pour la planification de leur enseignement en disent relativement peu sur les contenus lexicaux à aborder en classe. En France, par

exemple, la participation des notions lexicales à l'enseignement de la langue demeure plus que symbolique. Si nous mettons l'accent sur l'absence de notions fondamentales structurant l'enseignement du lexique, c'est parce que nombre de ces notions absentes sont en même temps celles sur lesquelles repose la véritable compréhension de notions grammaticales fondamentales, présentées dans les programmes et dans la pratique de l'enseignement comme autonomes de la réalité lexicale de la langue. Pour ne citer ici qu'un exemple de ce fait : comment est-il possible de faire véritablement assimiler le concept de circonstant sans le contraster avec celui d'actant, et comment ce dernier peut-il être assimilé sans que soit assimilé, de façon explicite ou non, le concept de structure prédicative des unités lexicales?

Legrand (2004) observe également que les préoccupations lexicales sont peu présentes dans les programmes officiels français, comparativement à d'autres composantes de la langue. La publication des dernières *Instructions Officielles* (I.O.) de l'école primaire française (2008) illustre néanmoins un certain regain d'intérêt envers le lexique, en réintroduisant par exemple un travail spécifique sur le lexique dès le cours préparatoire (CP).

Au Québec, le programme officiel de formation des élèves du primaire (MEQ, 2001) présente un certain nombre d'informations relatives au lexique : *Mots variés, corrects, précis, évocateurs liés aux thèmes abordés en français et dans les autres disciplines ; Termes utilisés pour consulter les outils de référence ; Vocabulaire visuel constitué des mots fréquents et utiles ; Forme de 3000 mots fréquents ; Familles de mots, Formation de mots (base, préfixe, suffixe), Expressions figées, expressions régionales, expressions courantes, Sens commun et sens figuré des mots*. Le programme gagnerait cependant à être plus complet et à mieux représenter, dans les éléments à l'étude, les composantes de la connaissance lexicale, de même que les interactions entre lexique, syntaxe et morphologie. Grossmann (2011) souligne par ailleurs la nécessité « [d'] établir une progression des notions de base nécessaires à la compréhension du système lexical et mieux réfléchir aux formes que peut prendre l'apprentissage 'incident', à partir de genres discursifs variés » (p. 164).

La didactique du lexique a mis au jour un certain nombre de notions fondamentales et les sources théoriques de référence sont suffisamment développées pour fournir le matériau de base nécessaire et suffisant pour établir la progression structurée dont parle Grossmann. Or, l'ensemble de ces notions de base est peut-être encore mal organisé et gagnerait à être systématisé davantage afin d'assurer une progression cohérente et systématique des notions aux fins d'une meilleure compréhension du système lexical par les enseignants et les élèves, et d'un meilleur enseignement/apprentissage du lexique. C'est en fait l'analyse des savoirs théoriques de référence qui entre ici en jeu. Jonnaert (1998) et Simard (1997) soulignent ainsi que la question qui se pose au didacticien d'une discipline est celle des référents qui doivent assier la formation didactique des enseignants (ce qui implique entre autres une part des contenus à enseigner aux élèves eux-mêmes). Hudson (2004) note à cet effet l'utilité de la linguistique pour l'enseignement :

As a discipline, linguistics is able to provide a framework into which to put the various language topics otherwise sometimes dealt with by teachers in a slightly haphazard way (Hudson, 2004 : 107).

Par conséquent, afin de déterminer les connaissances disciplinaires que l'enseignant doit maîtriser pour enseigner le lexique au primaire en français langue première (et, de façon parallèle, les contenus devant figurer dans les programmes), il faut d'abord faire appel à un modèle linguistique théorique du fonctionnement et de l'organisation du lexique, soit un modèle qui propose une description des savoirs lexicologiques. Nous sommes convaincus, à l'instar de Jonnaert et de Simard, qu'une telle réflexion a sa place dans toute didactique d'une discipline, et en particulier en didactique du français, où les savoirs de référence ne sont pas toujours aussi bien définis et systématisés qu'en didactique des mathématiques ou en didactique des sciences, par exemple.

### 2.3 Nécessité d'un fondement théorique des savoirs lexicologiques

Les didacticiens du lexique en français langue première, s'ils présentent un certain nombre de notions de base permettant d'aborder l'étude et l'enseignement du lexique (par exemple Galisson, 1979 ; Gosselin & Simard, 1995 ; Léon, 1998 ; Picoche, 1993 ; Simard, 1994), ne s'appuient pas sur une théorie lexicologique en particulier (sauf Picoche qui s'inspire de la linguistique de Gustave Guillaume). Quant aux théories linguistiques privilégiant la description théorique du lexique (Cruse, 1986 ; Pustejovsky, 1995), elles ne proposent pas toutes un appareil formel permettant de décrire l'ensemble des phénomènes lexicaux ou encore les composantes de la compétence lexicale et leurs interactions avec les autres modules linguistiques. De plus, elles n'ont pas nécessairement donné lieu à des applications pédagogiques ou à des démarches de didactisation des savoirs, et elles ne proposent pas non plus de descriptions des unités lexicales suffisamment nombreuses pour être utilisées comme support pédagogique (en d'autres mots, elles ne produisent pas de dictionnaire). Ces théories ne sont donc pas les plus appropriées pour servir de référence à la description des savoirs lexicologiques à des fins didactiques<sup>1</sup>. Enfin, bon nombre de publications en anglais (L1 et L2) dans le champ de la « linguistique appliquée » (*applied linguistics*) font appel à une vision lexicale du fonctionnement de la langue (Lewis, 1993 ; Nation, 2001 ; Schmitt, 2000, pour ne citer que ceux-là), mais sans forcément présenter les notions lexicales de façon formelle, systématique et structurée, ni en s'appuyant sur un modèle linguistique théorique spécifique.

La théorie Sens-Texte, plus particulièrement la branche lexicologique de celle-ci, la Lexicologie Explicative et Combinatoire (Mel'čuk et collab., 1995), propose quant à elle un système notionnel solide qui présente de façon formelle l'ensemble des savoirs lexicologiques (et le métalangage associé) permettant d'aborder l'étude du lexique. L'intérêt de cette théorie lexicologique est qu'elle comporte un système conceptuel riche et organisé permettant de décrire l'ensemble des phénomènes lexicaux, quelle que soit la langue. Elle met également clairement en évidence les rapports étroits qu'entretient le lexique avec la grammaire – notamment, avec la syntaxe et la morphologie. Ceci s'explique entre autres par le fait que la théorie Sens-Texte propose un modèle global du fonctionnement de la langue (sémantique, syntaxe, morphologie, phonologie). En outre, cette théorie a déjà fait l'objet de nombreuses applications dans le domaine de l'enseignement/apprentissage de la langue, notamment en français, en anglais et en russe (Gentilhomme, 1992 ; Leed & Nakhimovsky, 1990 ; Mel'čuk, 1992 ; Polguère, 2000 ; Steele, 1990 ; Tremblay, 2009), ce qui démontre l'intérêt qu'elle représente d'un point de vue didactique. Elle encadre enfin la production de dictionnaires (notamment d'un dictionnaire pédagogique, *Le Lexique Actif du Français* (Mel'čuk et Polguère, 2007)) et de bases de données lexicales, ressources indispensables pour l'enseignement et l'apprentissage du lexique.

Les savoirs lexicologiques à la base du système théorique et descriptif de la Lexicologie Explicative et Combinatoire constituent un ensemble notionnel riche et bien structuré. Nous croyons que les futurs enseignants de français au primaire et au secondaire peuvent en acquérir les notions de base, d'autant plus qu'une bonne partie de ces notions font déjà partie des programmes de formation québécois (primaire et secondaire)<sup>2</sup>, sans toutefois y être présentées de façon structurée. L'acquisition de cet ensemble structuré de notions favoriserait selon nous une meilleure compréhension du contenu des programmes et du système lexical lui-même, et contribuerait certainement à en améliorer l'enseignement (dépasser les listes de vocabulaire et adopter une perspective un peu plus systémique). Ces notions apparaissent dans un ouvrage d'introduction à la sémantique et à la lexicologie (Polguère, 2008). Or, bien que l'ensemble des savoirs lexicologiques fassent l'objet d'une caractérisation poussée à l'intérieur de cet ouvrage, on n'en connaissait pas, jusqu'à tout récemment, de modélisation formelle et ontologique. Une telle modélisation des connaissances est appropriée pour mettre en évidence l'organisation des savoirs, ce qui permet d'en faciliter la didactisation (Tremblay & Lefrançois, soumis) et, ultimement, d'en favoriser l'apprentissage.

### 3 Vers une ontologie des savoirs linguistiques... et lexicologiques

Le document numérique ayant maintenant supplanté le document imprimé, les connaissances nécessitent un traitement informatique pour être partagées et réutilisées. Les ontologies informatisées répondent

précisément à cette demande, car elles permettent de spécifier l'ensemble des concepts relevant d'un domaine donné et d'encoder formellement les liens existants entre ces concepts. En effet, une ontologie est une spécification explicite d'une conceptualisation (Gruber, 1993), en d'autres mots, une description (informatisée) des concepts d'un domaine et des relations qui peuvent exister entre eux. Le caractère formel de la description des concepts rend l'ordinateur capable de traiter les connaissances ainsi « ontologisées ». On considère de plus que les ontologies sont pertinentes pour la systématisation des connaissances et la planification de l'enseignement dans des contextes de formation à distance (Mizoguchi & Bourdeau, 2002 ; Psyché, 2007). Nous œuvrons donc à modéliser les savoirs lexicologiques de la Lexicologie Explicative et Combinatoire à l'intérieur d'une ontologie, que nous allons présenter.

### 3.1 Le Gros Tas de Notions (GTN) : une ontologie des notions linguistiques

Le travail de construction d'une ontologie des savoirs lexicologiques dont il va être question s'inscrit dans le problème plus général de la description et de la structuration logique des notions linguistiques fondamentales mises en jeu dans les activités métalinguistiques telles que la recherche linguistique elle-même et l'enseignement de la langue. Ce projet – ancré dans les travaux de lexicologie, lexicographie et terminologie de l'Observatoire de linguistique Sens-Texte (OLST)<sup>3</sup> –, s'est donc jusqu'à présent concentré sur les notions liées au lexique et à la sémantique lexicale, telles que celles présentées dans Polguère (2008). Cependant, la modélisation de ce type de notions mène nécessairement à la prise en compte de notions grammaticales incontournables. Il s'agit de celles relevant principalement de l'interface sémantique-syntaxe (prédicat, valence, actant, dépendance syntaxique, syntagme, régime syntaxique, etc.) et de la morphologie (mot-forme, morphe, morphème, flexion, dérivation, etc.).

Nous allons présenter l'ontologie des notions de linguistique appelée par autodérision le *Gros Tas de Notions*, désormais GTN. Ce modèle ontologique (qui est toujours en construction) a été développé à l'aide de l'éditeur d'ontologies Protégé (conçu à l'Université Stanford<sup>4</sup>). Nous avons suivi les étapes de construction d'une ontologie Protégé présentées dans Natalya & McGuinness (2001). La prochaine section présente un bref survol des ontologies du domaine déjà existantes. Nous nous concentrons ensuite sur la description des caractéristiques de notre ontologie qui présentent un intérêt particulier pour le domaine de la didactique du lexique.

### 3.2 Ontologies existantes

Il existe deux grands types d'ontologies dans le domaine linguistique : les ontologies linguistiques et les ontologies métalinguistiques. Il est nécessaire d'examiner ces ontologies, dans la mesure où elles proposent des descriptions formelles des notions de référence d'un domaine (ici la linguistique) et illustrent ainsi la cohérence du système représenté. La cohérence du système de référence constitue un des critères que mettent en évidence Boivin & Pinsonneault (2008) en vue de garantir la validité de la démarche de didactisation.

#### 3.2.1 Ontologies linguistiques

Les ontologies linguistiques sont les plus courantes. Il s'agit d'ontologies servant à décrire le vocabulaire d'une langue. Elles sont plus particulièrement destinées à être utilisées dans une perspective de Traitement Automatique de la Langue (TAL). Le WordNet de l'Université Princeton (Felbaum, 1997) en est sans doute la plus connue des ontologies linguistiques. Il s'agit d'une base de données lexicales de l'anglais<sup>5</sup> (appelée ontologie à cause de la structure hiérarchique à laquelle elle répond), qui comprend différents types d'entités lexicales (mots-composés, collocations, locutions), mais dont l'unité lexicale constitue l'unité linguistique principale de description. Sa structure n'est cependant pas organisée autour des unités lexicales individuelles. En effet, la base de données est organisée principalement à partir de la relation d'hyper-/hyponymie connectant non les unités lexicales, mais des regroupements d'unités lexicales synonymes, appelés *synsets*. Chaque synset fonctionne comme unité de structuration de

l'ontologie, et représente plus qu'un mot lui-même ; dans une perspective cognitive, le synset s'approche d'une unité psycholinguistique de raisonnement. En ce sens, WordNet peut bel et bien être considéré comme une ontologie, dans la mesure où des concepts (et pas seulement des « mots ») y sont explicitement représentés.

En terminant, notons, à l'instar de Rastier (2004), que les oppositions ontologiques servant à orienter la description linguistique vont conduire le concepteur à privilégier l'être au faire (parce que plus caractéristique de « l'être en tant qu'être » au sens aristotélicien) et placer ainsi les noms au cœur de la description linguistique – ce qui est effectivement le cas dans les ontologies actuelles comme WordNet. Cette caractéristique est probablement une des lacunes des ontologies linguistiques, qui basent la représentation et la structuration lexicale sur des unités nominales.

### 3.2.2 Ontologies métalinguistiques

Les ontologies du second type, que l'on pourrait qualifier de « métalinguistiques », ne décrivent pas exactement les langues elles-mêmes, mais en précisent les concepts descriptifs, concepts qui rendent compte de la structure de la langue à l'un ou l'autre de ses niveaux de fonctionnement (lexical, syntaxique, morphologique, phonologique). Certaines des ontologies métalinguistiques existantes se rattachent à des ontologies dites « supérieures ». C'est le cas de l'ontologie GOLD (*General Ontology for Linguistic Description*)<sup>6</sup>, la première ontologie construite spécifiquement pour la description linguistique sur le Web sémantique. GOLD se rattache à une ontologie de type supérieur, SUMO (Farrar & Langendoen, 2003), qui fait état d'un certain nombre de catégories, en distinguant par exemple les concepts « physique » et « abstrait » de même que les concepts « objet » et « processus ». Elle fournit des axiomes qui permettent de régulariser le comportement et les propriétés de ces entités sous la forme de règles d'inférence. Les concepteurs de GOLD soulignent que ce qui justifie le classement des catégories linguistiques n'est pas le comportement morphosyntaxique des expressions, lequel varie d'une langue à l'autre, mais bien le type d'entité ontologique que ces catégories peuvent représenter :

The criteria we used for grouping grammatical categories are based, not on the morphosyntactic behavior of expressions which varies widely from language to language, but on what kind of ontological entities they represent (Farrar et Langendoen, 2003 : 3).

Différents types d'entités sont ainsi identifiés comme des expressions linguistiques. Par exemple, des catégories nécessaires pour la description linguistique, comme **mot**, **morphème**, etc. sont placées au même niveau qu'une catégorie comme **langage**. Or, ce classement de catégories linguistiques à partir de catégories ontologiques prédéfinies ne fonctionne pas nécessairement très bien. Par exemple, sur le plan morphologique, les langues manifestent des comportements souvent extrêmement différents et ne peuvent être classifiées de la même façon à l'intérieur d'une hiérarchie d'entités ontologiques. Un autre problème se pose quant à l'ontologie GOLD : les concepts sous-jacents relèvent pratiquement uniquement du domaine de la morphologie et de la syntaxe alors que d'autres axes descriptifs essentiels, ceux servant à décrire le lexique par exemple, sont négligés.

Les développements récents dans le domaine des ontologies pour la description des langues, et en particulier des lexiques, ont malgré tout montré qu'il était possible de créer des ontologies jouant le rôle de modèle pour le traitement des ressources lexicales. C'est le cas de la norme ISO 24613:2008 appelée *Lexical Markup Framework* ou LMF (*cadre de balisage lexical*, en français). Il s'agit d'un modèle qui fournit un cadre standard commun pour le traitement automatique des lexiques et ressources lexicales (Francopoulo et coll., 2006). Le cadre de modélisation de LMF a pour but de faire l'inventaire des composantes essentielles d'un modèle lexicographique générique (celui-ci étant indépendant d'une théorie lexicographique particulière), de décrire les contraintes qui régissent leur agencement, de même que d'inventorier les descripteurs qui leur sont associés. Pour y parvenir, les auteurs de la norme LMF (soit une soixantaine de collaborateurs) ont d'abord commencé par rassembler les descriptions de dictionnaires connus (y compris les bases de données lexicales comme WordNet et des modèles lexicaux issus de la Lexicologie Explicative et Combinatoire, qui est notre cadre de référence pour le GTN) afin

d'établir une terminologie consensuelle permettant de décrire les composantes des lexiques décrits dans ces dictionnaires et bases de données lexicales. Ensuite, un modèle a été conçu pour représenter ces dictionnaires en détail. Dans l'élaboration de ce modèle, tous les niveaux de fonctionnement des énoncés (sémantique, syntaxe, etc.) ont été pris en compte, en plus du lexique. La morphologie a reçu une attention spécifique afin qu'il soit possible d'élaborer des dispositifs adéquats permettant de traiter les langues présentant des difficultés sur le plan du traitement morphologique. Le modèle développé après plusieurs années de travail sert de base à la normalisation des ressources lexicales en vue d'assurer leur interopérabilité<sup>7</sup>.

LMF représente une ressource intéressante, mais la description qu'il propose est en quelque sorte un « meltingpot » théorique et ne poursuit pas de visées pédagogiques. En fait, à notre connaissance, aucune des ontologies du domaine de la linguistique n'a donné lieu à des applications en didactique des langues et en enseignement/apprentissage du français ou à des démarches de didactisation, malgré le potentiel des ontologies dans une perspective de formation, comme le montrent certaines recherches en éducation (Desmoulin, 2010 ; Desmoulin & Grandbastien, 2006 ; Psyché, 2007)

L'ontologie GTN, qui a été élaborée à l'OLST entre 2006 et 2009, et qui est toujours en développement aujourd'hui, possède quant à elle un certain nombre de caractéristiques qui présentent un intérêt certain pour la didactique du lexique et pour la didactique des langues en général. Ces caractéristiques vont maintenant être introduites.

### 3.3 Caractéristiques du GTN

Nous allons détailler trois caractéristiques de l'ontologie GTN, qui nous semblent pertinentes d'un point de vue didactique : 1) la distinction fondamentale entre terme, concept et notion, qui illustre entre autres toute l'importance du métalangage dans une discipline ; 2) la structuration hiérarchique à héritage multiple, qui permet de mieux comprendre comment les concepts linguistiques de référence de la didactique du français (et de la didactique du lexique en particulier) sont organisés logiquement et comment ils se comprennent les uns par rapport aux autres, selon le niveau hiérarchique auquel ils apparaissent ; 3) la structuration sous forme de réseau, qui permet d'illustrer les liens définitionnels entre les concepts et qui permet par la suite d'organiser un enseignement cohérent et méthodique des notions et, potentiellement de favoriser l'apprentissage. Comme nous le verrons, ces caractéristiques sont liées aux critères de Boivin & Pinsonneault (2008) relatifs à la didactisation des notions linguistiques de référence de la didactique du français.

#### 3.3.1 Distinction entre terme, concept et notion

Cette distinction met en évidence l'existence de trois entités informationnelles : les termes, les concepts et les notions. Les termes et les concepts sont représentés dans le GTN par des classes différentes, ce qui est unique dans le domaine de la modélisation ontologique. Dans les ontologies courantes, un terme n'est qu'une étiquette apposée sur un concept. Dans notre modélisation, les termes relèvent d'une classe particulière et possèdent tout un ensemble de caractéristiques. C'est l'association entre le terme et le concept qui fonde l'existence d'une notion. Cette façon de distinguer explicitement concept, terme et notion est ainsi tout à fait originale. On trouvera une présentation plus détaillée de ces notions dans Polguère (à paraître).

De façon plus spécifique, nous appelons *concept* une **entité informationnelle** pouvant être impliquée dans un raisonnement quelconque, c'est-à-dire dans une inférence. L'ensemble des concepts est par nature infini et non dénombrable. De plus, un concept – c'est-à-dire une entité informationnelle impliquée dans une inférence – n'est pas nécessairement « nommable ». Nous effectuons chaque jour des milliards d'inférences et il est vraisemblable qu'une grande partie des concepts impliqués dans ces inférences ne peuvent être désignés au moyen d'une unité lexicale de notre langue maternelle. Ainsi, un ouvrage comme le génial *Baleinié*, ou *Dictionnaire des tracas*, de C. Murillo, J.-C. Leguay et G. Østermann (Éditions du Seuil), satisfait un besoin bien légitime en nous offrant des unités lexicales permettant de

verbaliser des tracas de la vie courante que nous avons déjà clairement conceptualisés – par exemple, dans le tome 2 (page 127), le verbe de sentiment SIRMOUFFER, défini par 'sentir que la personne arrivée après vous va essayer de passer devant'. Dans l'ontologie, les concepts sont caractérisés par différentes propriétés, par exemple le domaine duquel relève le concept décrit (par exemple, le concept de polysémie est un concept qui relève du domaine de la lexicologie et de la sémantique), les concepts définitionnels qui y sont associés (les concepts définitionnels du concept de polysémie sont entre autres les concepts de sens et de flexion), le terme qui le désigne en français et en anglais, etc. En son état actuel, l'ontologie GTN contient plus de 200 concepts.

Nous désignons par **terme** une **unité lexicale** qui dénote un concept relevant d'une discipline donnée, et plus spécifiquement ici, de la linguistique. En tant qu'unité lexicale, le terme appartient de fait au lexique de la langue : il peut recevoir une définition lexicale, il entretient des liens multiples avec d'autres unités lexicales (hyperonymie, synonymie, antonymie, dérivation, etc.) et il possède ses propriétés de combinatoire lexicale et grammaticale. On notera aussi que le terme appartient nécessairement à une langue donnée, et peut donc être traduit, alors que le concept est une entité de niveau plus profond que le niveau linguistique. Dans l'ontologie, la classe qui contient les termes permet d'encoder les différentes informations associées à un terme donné : sa définition lexicographique, le terme équivalent en anglais (ou dans toute autre langue, ce qui rend compte du caractère multilingue de notre ontologie), les synonymes de ce terme s'il y a lieu (*phrasème complet* pour *locution*, par exemple), ou encore ses dérivés sémantiques (sous le terme *grammatical*, on trouvera le terme *grammaticalité*). L'ontologie contient ainsi davantage de termes que de concepts, soit environ 320 termes anglais et 330 termes français.

Finalement, nous appelons **notion** une **association** entre un concept et un terme permettant de désigner ce concept dans des énoncés écrits dans une langue donnée. Dans la pratique scientifique, nous ne manipulons véritablement que des notions, puisqu'il nous est impossible de parler scientifiquement de concepts sans proposer des termes spécifiques pour les désigner. Deux notions ne pourront être considérées comme identiques que si 1) elles renvoient exactement au même concept et 2) elles mettent en jeu deux termes qui sont des synonymes ou des traductions exactes. Ainsi, dans la terminologie sur laquelle nous nous appuyons (Mel'čuk et coll. 1995; Polguère 2008), la notion d'**unité lexicale** est strictement identique à celles de **lexie** et de **lexical unit** en anglais, puisque, dans notre terminologie, *lexie* est un synonyme exact d'*unité lexicale* et *lexical unit* en est sa traduction anglaise, les trois termes renvoyant au même concept. L'ontologie comprend pour le moment plus de 200 notions, dont une centaine étiquetées comme des notions relevant de la lexicologie.

La distinction établie dans l'ontologie entre concept, terme et notion nous permet de rappeler à quel point la maîtrise du métalangage est fondamentale dans toute discipline, ce à quoi n'échappe pas l'enseignement du français. Pourtant, des recherches (par exemple Perrault (2006)) montrent que les enseignants maîtrisent mal le métalangage utilisé pour la description du français, et de la grammaire en particulier. Peut-être la difficulté à maîtriser ce métalangage s'explique-t-elle par le fait que les notions ne sont pas suffisamment bien définies, ou ont des définitions variées selon les cadres théoriques de référence, ce qui explique d'ailleurs qu'un groupe d'uniformisation de la terminologie grammaticale ait été créé il y a quelques années pour remédier à ce problème (Chartrand, Lord & Gauvin, 2010). Dans l'ontologie GTN, le fait d'associer explicitement un terme donné à un concept précis permet la mise en valeur d'une terminologie unifiée.

### 3.3.2 Structuration hiérarchique

La structure typique d'une ontologie correspond à une architecture en arborescence qui repose formellement sur le lien ontologique « is a ». Ce lien connecte un concept B à un concept A, le concept B étant un « type particulier » de A. Précisons que la structuration hiérarchique n'est pas une caractéristique du GTN lui-même, puisqu'une ontologie est par nature hiérarchiquement organisée<sup>8</sup>, mais cette caractéristique est importante d'un point de vue didactique, raison pour laquelle nous en traitons ici.

Le travail de conceptualisation que nous avons effectué afin d'organiser les notions de lexicologie dans cette structure hiérarchique nous a conduits à créer de nouvelles classes conceptuelles afin de pouvoir



répartir logiquement les concepts à travers différents types de classes. Par exemple, afin de caractériser le concept de polysémie, nous avons créé une classe conceptuelle qui n'existait pas formellement auparavant dans le système notionnel de référence de la Lexicologie Explicative et Combinatoire. Il s'agit de la classe LINGUISTIC\_PROPERTY<sup>9</sup> (propriété linguistique). Cela nous a permis pas la suite de « ranger » dans cette classe non seulement le concept de polysémie, mais d'autres concepts qui représentent des propriétés linguistiques, par exemple le concept de grammaticalité, qui renvoie à la propriété d'un énoncé linguistique d'être conforme aux règles grammaticales d'une langue L, ou encore le concept de prédicativité, qui constitue une propriété du sens des unités lexicales.

L'intérêt d'une telle classification, d'un point de vue didactique, est que les classes en question peuvent être présentées telles quelles ou en partie pour favoriser l'apprentissage de certaines notions. Par exemple, comprendre qu'il existe des propriétés qui s'appliquent à différents types d'entités linguistiques (phrase, unité lexicale, morphème, etc.) donne un point de vue global sur le fonctionnement de la langue et permet, tout en décloisonnant les matières (sémantique, syntaxe, morphologie,).

L'autre aspect intéressant qui découle de l'organisation hiérarchique est que les concepts de même niveau partagent un certain nombre de caractéristiques communes, ce qui permet d'envisager qu'on les présente de façon parallèle. Dans le module de cours en didactique du lexique qui a été développé à partir de l'ontologie, certaines notions sont ainsi introduites en parallèle, par exemple les notions *vocabulaire [d'une langue]*, *vocabulaire [d'un individu]* et *lexique [d'une langue]*, représentées dans la Figure 1 ci-dessous:

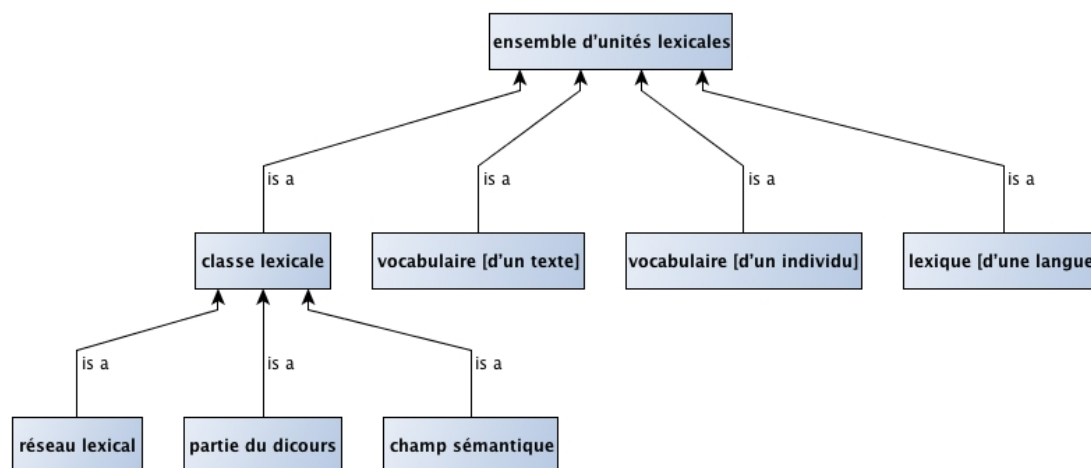


Figure 1 : Extrait du réseau notionnel contrôlé par la notion *ensemble d'unités lexicales*

Dans le réseau de la Figure 1, les liens hiérarchiques sont représentés par le lien « is a » ('est un').

Si la classification hiérarchique propre aux ontologies est puissante, elle n'épuise pas tous les types de relations qui peuvent exister entre les concepts d'un domaine. En effet, d'autres types de connexions existent entre les concepts, qui sont de nature plus transversale, comme nous allons maintenant l'illustrer.

### 3.3.3 Aspect réseau de l'ontologie

Tel que mentionné dans la section 3.3.1, le fait d'avoir distingué terme et concept dans l'ontologie nous a permis de séparer les informations relatives à chacun. Rappelons, par exemple, que les synonymes et les variantes morphologiques sont des éléments associés à un terme et non à un concept. Une définition est également une information que l'on associe au terme, bien que cette définition nous informe sur des caractéristiques du concept associé. Le travail de définition des termes a été central dans notre démarche de représentation ontologique : grâce à une démarche précise de définition des termes (définition par

genre prochain et différence spécifique), nous avons pu identifier les concepts « définitoires » associés à un concept donné. Cette spécification des liens entre les concepts contribue à faire de l'ontologie un immense réseau où tous les concepts sont liés.

La navigation dans l'ontologie peut donc se faire à partir de la structuration hiérarchique, tel que nous l'avons illustré dans la section précédente, mais aussi à travers les liens définitionnels qui connectent les concepts entre eux. Le moteur de recherche associé à l'ontologie permet d'effectuer des requêtes sur ces liens. Ces requêtes peuvent également être visualisées sous forme de graphes, qui permettent d'explicitier les liens entre concepts dans l'ontologie. La Figure 2 en est une illustration.

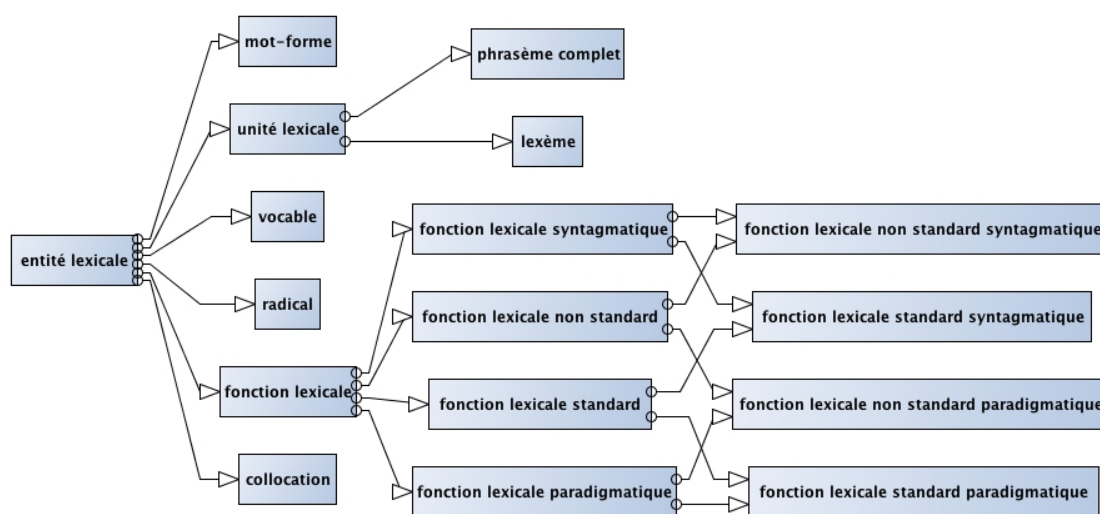


Figure 2 : Exemple de représentation graphique

Lors de l'élaboration d'un module de cours en didactique du lexique (Tremblay, 2009), nous avons utilisé le moteur de recherche de Protégé afin de naviguer dans le réseau de concepts et d'organiser de façon parallèle les contenus du cours. Cette façon originale de sélectionner et de structurer les éléments de contenus d'un cours répond au besoin de faire appel à une démarche rigoureuse de didactisation des savoirs de référence tout en rendant transparent une partie du processus de transposition didactique lui-même (Chevallard, 1985/1991). Le choix des notions à examiner et les parcours subséquents dans les réseaux sont en effet systématiquement présentés et justifiés, jusqu'à la sélection et l'organisation finales des contenus du module de cours en question.

Nous avons ainsi commencé par examiner la notion *lexique*, bien évidemment centrale dans un cours universitaire de didactique du lexique, et dans l'enseignement du lexique auprès des élèves. Cet examen nous a ensuite conduit à la notion *unité lexicale* et à la notion *lexème* (voir Figure 4), l'unité lexicale (lexème ou locution) étant l'objet d'étude premier de l'enseignement/apprentissage du lexique.

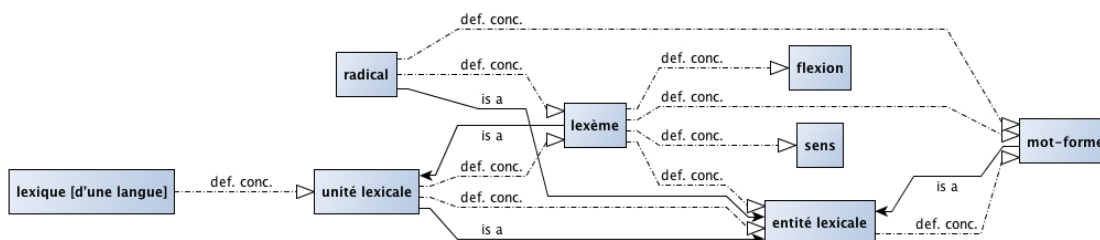


Figure 3 : Réseau de notions pour *lexème*

Afin de « lire » le réseau de la figure précédente, il faut partir de la notion *lexème* pour retrouver ensuite les notions définitionnelles (identifiées par le lien **def. concept**) vers lesquelles cette notion pointe (*entité lexicale*, *sens*, *mot-forme* et *flexion*). Rappelons que ces notions apparaissent dans la définition du terme *lexème* : *entité lexicale qui est structurée autour d'un sens et qui est exprimable par un ensemble de mots-formes que seule distingue la flexion*.

L'intérêt des réseaux, du point de vue didactique, est qu'ils nous fournissent des pistes sur la façon de présenter les notions. Par exemple, pour la notion *lexème*, puisque le lexème est d'un type d'entité lexicale, cette notion sera mise en contraste avec les autres notions qui dénotent des types d'entités lexicales, par exemple *mot-forme* et *locution*. La notion de mot-forme permet de distinguer les différents types d'entités lexicales désignées par le terme *mot*, notamment les notions *lexème* et *locution*, présentées dans le réseau notionnel de la Figure 4.

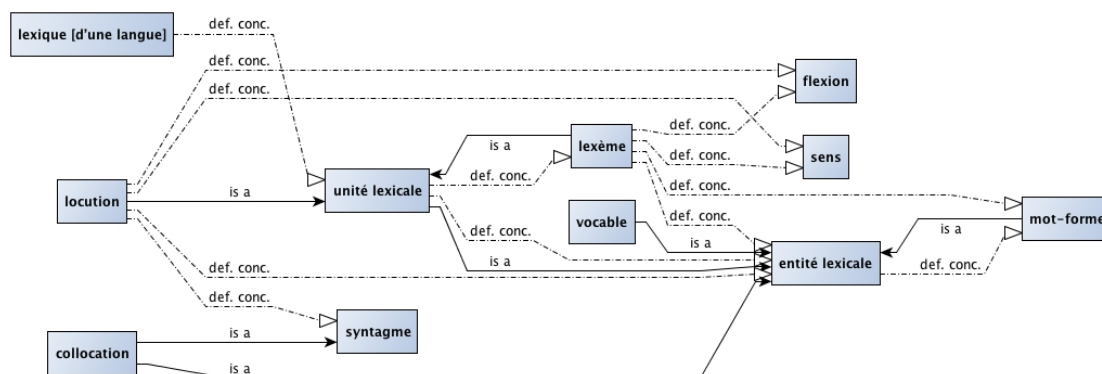


Figure 4 : Réseau de notions pour *locution* et *lexème*

Les notions définitionnelles partagées par *lexème* et *locution*, en tant que types d'unités lexicales, sont les suivantes : *entité lexicale*, *sens* et *flexion*. Ce qui les distingue, c'est principalement le fait qu'une locution est constituée d'un ensemble de syntagmes, comme l'illustre la définition de locution : *entité lexicale qui est structurée autour d'un sens et qui est exprimable par un ensemble de syntagmes que seule distingue la flexion*, alors qu'un lexème est constitué d'un ensemble de mots-formes.

Ces exemples de réseaux illustrent de quelle façon il est possible d'introduire, en contexte de formation à l'enseignement, les notions qui permettent de mieux comprendre les propriétés des unités qui composent le lexique, de même que leur fonctionnement. Nous abordons maintenant les applications pédagogiques possibles du GTN.

#### 4 Utilisations pédagogiques du GTN

La création d'une ontologie comme le GTN demande de nombreuses heures de travail, d'abord pour se familiariser avec les outils de représentation des connaissances que sont les éditeurs d'ontologie et ensuite pour modéliser l'ensemble des concepts du domaine à représenter. En ce sens, l'élaboration d'une ontologie n'est pas un projet didactique à proprement parler ; néanmoins la ressource qu'elle constitue fournit un cadre très pertinent pour la sélection, l'organisation et la structuration des contenus linguistiques de référence de la didactique du lexique. C'est ce que nous avons pu vérifier en nous servant de l'ontologie pour appuyer l'élaboration d'un module de cours en didactique du lexique destiné à de futurs maîtres au primaire au Québec (Tremblay, 2009).

À partir d'un parcours logique dans les réseaux de l'ontologie (tel qu'illustré à la section 3.3.3), nous avons pu organiser le contenu de ce module de cours et répartir les notions lexicales à travers cinq séances de cours, l'équivalent d'un module de cours universitaire de 15 heures. Nous y avons ajouté des

pistes de travail concrètes pour l’enseignement du vocabulaire en classe, que nous ne présentons pas ici, faute d’espace. Le Tableau 1 présente cette progression.

Séance 1	lexique [d’une langue], vocabulaire [d’un texte], vocabulaire [d’un individu], unité lexicale, lexème, locution, mot-forme, sens, flexion, syntagme, vocable, polysémie/monosémie,
Séance 2	métaphore, métonymie, analogie, proximité, intersection sémantique, dérivation sémantique, synonymie, antonymie, généralité (hyperonymie, hyponymie), conversivité, prédicat sémantique, objet sémantique, actant sémantique
Séance 3	définition, définition analytique, genre prochain, différences spécifiques, champ sémantique
Séance 4	collocation, base d’une collocation, collocatif, combinatoire, combinatoire restreinte, régime
Séance 5	dictionnaire

Tableau 1 : Progression des contenus d’un module de cours en didactique du lexique

Une telle progression répond en partie à la remarque de Grossmann (2011) quant à la nécessité d’établir les notions de base nécessaires à la compréhension du système lexical. Cette progression, établie selon une démarche rigoureuse à partir d’une modélisation formelle des savoirs lexicologiques nous semble un apport pertinent à la didactique du lexique, du point de vue des fondements linguistiques de cette branche de la didactique du français. La progression établie est également en conformité avec les contenus relatifs au lexique qui sont présentés dans les programmes québécois de formation des élèves, notamment dans la section « Lexique » de la *Progression des apprentissages* (MELS, 2009), tout en proposant une perspective plus structurée quant à l’ordre selon lequel présenter les notions à enseigner. Enfin, cette organisation des notions lexicales de base de l’enseignement du vocabulaire sert de fondement à l’élaboration d’un manuel d’enseignement/apprentissage du vocabulaire destiné aux enseignants du primaire (Tremblay & Ancil, en préparation), ce qui illustre l’apport de l’ontologie dans le champ de la didactique du français.

De façon plus générale, le fait de pouvoir recourir à une modélisation formelle telle que celle présentée dans le GTN permet de guider de manière méthodique le processus d’élaboration d’un cours portant sur des notions linguistiques, comme c’est le cas en didactique du français ou en enseignement du français. Par exemple, même si un enseignant sait souvent de façon intuitive, ou par expérience, qu’une notion pose davantage problème qu’une autre aux étudiants ou aux élèves (parce que cette notion est plus complexe que d’autres, notamment), cette intuition ne repose pas forcément sur une justification didactique en lien avec le système notionnel de référence duquel relève la notion en question. Dans cette optique, l’originalité de l’ontologie est de supporter la structuration des contenus en permettant de mettre en évidence : 1) le « poids » d’une notion dans le système (par exemple, la notion d’unité lexicale est une des notions les plus fréquentes dans l’ontologie, car elle est définitionnelle pour un grand nombre de notions linguistiques : cette information nous indique que la notion d’unité lexicale est une notion incontournable, qui devrait figurer parmi les premières notions d’un cours sur le lexique ou le vocabulaire, quel que soit le niveau scolaire envisagé) ; 2) la complexité conceptuelle des notions (par exemple, la notion de collocation est une des notions lexicales qui comprend le plus de notions définitionnelles. Si l’on veut faire l’étude de la notion de collocation, d’ailleurs au carrefour de la sémantique et de la syntaxe, on doit donc d’abord maîtriser les notions qui la définissent et la caractérisent : *unité lexicale, syntagme, locuteur, sens, combinatoire restreinte, base [d’une collocation]* et *collocatif*). Ces deux derniers points constituent à notre avis des exemples assez illustratifs de l’utilisation de l’ontologie pour la justification du choix des notions à enseigner et de leur progression dans un syllabus de cours.

Enfin, dans la perspective d'une réflexion sur les contenus linguistiques des programmes de formation (tant au Québec qu'en France), l'intérêt de l'ontologie réside dans deux aspects. Premièrement, l'ontologie unifie les notions grammaticales et lexicales, en explicitant les liens logiques qui les unissent. Deuxièmement, elle permet de gérer les équivalences terminologiques, autorisant ainsi le passage d'une terminologie « scientifique » à une terminologie « appliquée », et donc la mise en place d'une véritable réflexion sur le rapport entre terminologie linguistique et terminologie propre à la didactique de la langue, un travail déjà amorcé en didactique de la grammaire, tel que mentionné plus tôt. Le travail de réflexion et d'organisation de la terminologie scolaire demande à notre avis à être approfondi, à la manière de ce qu'a proposé Desmoulins (2010) dans son projet collaboratif (experts de la modélisation ontologique et enseignants) visant l'élaboration d'une ontologie pour l'enseignement des compétences en géométrie. Une telle démarche illustre d'autres possibilités d'utilisation de l'ontologie à des fins didactiques. Nous projetons par ailleurs la mise en place de recherches en milieu scolaire (au primaire et au secondaire), afin de mettre à l'épreuve et de valider la progression notionnelle proposée et de mesurer son impact sur le développement de la compétence lexicale des élèves.

## 5 Conclusion

D'un point de vue à la fois linguistique et didactique, le travail de conceptualisation inhérent à la construction d'une ontologie de notions linguistiques est important parce qu'il permet : 1) de vérifier la cohérence du système de référence (critère évoqué par Boivin & Pinsonneault, 2008 en lien avec la démarche de didactisation des notions linguistiques) ; 2) d'améliorer le système de notions d'une discipline, en classant ces notions à travers des classes hiérarchisées (par exemple : « système de règles linguistiques », « propriétés linguistiques », etc.) ; 3) d'affiner notre propre compréhension des notions linguistiques que nous enseignons. Cela nous paraît un défi intéressant autant pour les didacticiens du français qui réfléchissent aux fondements linguistiques de leur discipline que pour les linguistes intéressés aux questions d'enseignement et d'apprentissage de la langue. Cette démarche de représentation systématique des connaissances constitue une sorte de mise à l'épreuve du système de référence linguistique à la base de toute transposition didactique.

De plus, le fait que l'ontologie illustre clairement les liens entre les notions des différents modules de fonctionnement de la langue (lexique, syntaxe, morphologie) est un avantage certain pour quiconque s'intéresse à la question d'un enseignement « intégré » du français et des savoirs linguistiques de base traditionnellement transmis par l'école. Cela entraîne des retombées pratiques pour la formation des enseignants et pour celle de leurs futurs élèves, pour la spécification des contenus linguistiques de référence dans les programmes de formation des élèves et, plus généralement, pour le développement de la didactique du français en tant que discipline.

## Références bibliographiques

- Ancil, D. (2011). *L'erreur lexicale au secondaire : analyse d'erreurs lexicales d'élèves de 3e secondaire et description du rapport à l'erreur lexicale d'enseignants de français*. Thèse de doctorat non publiée, Montréal : Université de Montréal.
- Biemiller, A. (2011). Vocabulary development and implications for reading problems. In A. McGill-Franzen et R. L. Allington, *Handbook of reading disability research*, New York : Routledge, 208-218
- Boivin, M.-C. & Pinsonneault, R. (2008). Description grammaticale et transposition didactique : le cas des infinitives et des participiales. *Enjeux*, 71, printemps 2008, p. 29- 57.
- Carter, R. (1982). *Linguistics and the Teachers*. London/ Boston/ Melbourne/ Henly : Routledge & Kegan Paul.
- Chartrand, S.-G., Lord, M.-A. & Gauvin, I. (2010). La terminologie grammaticale pour l'enseignement du français langue première au Québec : état des travaux de l'équipe québécoise de l'AIRDF. *Lettre de l'Association*, 45-46, Namur : Belgique, 42-49.
- Chevallard, Y. (1985/1991). *La transposition didactique : du savoir savant au savoir enseigné*. (2e ed.). Grenoble : Éd. La Pensée sauvage.

- Cruse, D. A. (1986). *Lexical semantics*. Cambridge : Cambridge University Press.
- David, J. (2000). Le lexique et son acquisition : aspects cognitifs et linguistiques. *Le Français aujourd'hui*, 131, 21-27.
- David, J., Paveau M-A. & Petit, G. (2000). Présentation. Construire les compétences lexicales. *Le Français aujourd'hui*, 131, 12-18.
- Desmoulins, C. (2010). Construction avec des enseignants d'une ontologie des compétences en géométrie, Geoskills. Colloque IC 2010, Nîmes.
- Desmoulins, C., & Grandbastien, M. (2006). Une ingénierie des EIAH fondée sur des ontologies. In M. Grandbastien (dir.), *Environnements Informatiques et Apprentissage Humain*. Paris: Hermès-Lavoisier, 161-180.
- Desrosiers, H. & Ducharme A. (2006). Commencer l'école du bon pied. Facteurs associés à l'acquisition du vocabulaire à la fin de la maternelle. In *Étude longitudinale du développement des enfants du Québec (ÉLDEQ 1998-2010)*, Québec : Institut de la statistique du Québec, 4, 1.
- Dreyfus, M. (2004). Pratiques et représentations de l'enseignement du lexique à l'école primaire. *9e colloque de l'Association Internationale pour la Recherche en Didactique du Français*, Québec : Université Laval.
- Farrar, S. & Langendoen, T. (2003). A linguistic ontology for the semantic web. *Glott International*, 7, Number 3, March 2003.
- Fellbaum, C. (1997). *WordNet : An Electronic Lexical Database*. Cambridge : MIT Press.
- Franco-poulo G., George M., Calzolari N., Monachini M., Bel N., Pet M., Soria C. (2006). *Lexical Markup Framework (LMF)*. LREC, Genoa
- Galisson, R. (1979). *Lexicologie et enseignement des langues (essais méthodologiques)*. Paris : Hachette.
- Gentilhomme Y. (1992). Panorama sur le Dictionnaire Explicatif et Combinatoire : retombées pédagogiques. *Actes du colloque Enseignement apprentissage du lexique en français langue étrangère*. Strasbourg : Université des Sciences humaines. 97-150.
- Gosselin, M. & Simard, C. (1995). *Des mots à la parole, enseigner le vocabulaire*. Québec : Éditions du CRAP.
- Grossmann, F. (2011). Didactique du lexique : état des lieux et nouvelles orientations. *Pratiques*, 149/150, 163-183.
- Gruber, T.R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition*, 5(2), 199-220.
- Hudson, R. (2004). Why education needs linguistics (and vice versa). *Journal of linguistic*, 40(1), 105-130
- Jonnaert, Ph. (1998). Quels référents pour une formation didactique des enseignants ? Réflexions et questions, In M. Sachot (dir.), *Le référentiel d'apprentissage et de formation : un outil didactique ?* Strasbourg : Publications du CRDP d'Alsace, 141-159.
- Leed, R. L. & Nakhimosky, A. (1990). Lexical Functions and Language Learning. In J. Steele (Ed.), *Meaning-Text Theory : Linguistics, Lexicography, and Implications*. Ottawa/London/Paris : University of Ottawa Press, 365-374.
- Leeman, D. (2000). Le vertige de l'infini ou de la difficulté de didactiser le lexique. *Le Français aujourd'hui*, 131, 42-52.
- Legrand, G. (2004). Les rubriques du vocabulaire dans l'enseignement élémentaire : étude sur deux siècles de textes ministériels. In E. Calaque, F. Grossmann, M.-A Paveau. et G. Petit (Coords.), *Didactique du lexique: contextes, démarches, supports*. Bruxelles : de Boeck.
- Léon, R. (1998). *Enseigner la grammaire et le vocabulaire à l'école*. Paris : Hachette.
- Lewis, M. (1993). *The Lexical Approach*. Hove : Language Teaching Publications.
- Meara, P. M. (1996). The Dimensions of Lexical Competence. In G. Grown, K. Malmkjaer & J. Williams (Eds), *Competence and Performance in Language Learning*. Cambridge : Cambridge University Press. 33-51.
- Mel'čuk, I. (1992). La phraséologie et son rôle dans l'enseignement/apprentissage d'une langue étrangère. *Études de linguistique appliquée*, 92, 82-113.

- Mel'čuk, I., Clas, A. et Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot, Coll. Universités francophones.
- Mel'čuk, I. et Polguère, A. (2007). *Le Lexique Actif du Français*. Louvain-la-Neuve : Duculot.
- Ministère de l'Éducation, du Loisir et du Sport. (2009). *La Progression des apprentissages en français au primaire*. Québec.
- Ministère de l'Éducation du Québec. (2001). *Programme de formation de l'école québécoise du niveau préscolaire et primaire*. Québec : MEQ.
- Ministère de l'Éducation nationale. (2008). *Bulletin officiel hors série*. N° 3 du 19 juin 2008.
- Mizoguchi R., & Bourdeau J. (2002). Theory-Aware Authoring Environment : Ontological Engineering Approach. Paper presented at the *Proceedings of the ICCE Workshop on Concepts and Ontologies in Web-based Educational Systems*, Technische Universiteit Eindhoven.
- Natalya, N. & McGuinness, D. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York, NY: Cambridge University Press.
- Pagani, L. S., Fitzpatrick, C., Belleau, L. & Janosz, M. (2011). Prédire la réussite scolaire des enfants en quatrième année à partir de leurs habiletés cognitives, comportementales et motrices à la maternelle. In *Étude longitudinale du développement des enfants du Québec (ÉLDEQ 1998-2010) – De la naissance à 10 ans*, Québec, Institut de la statistique du Québec, vol. 6, fascicule 1.
- Perrault, M. (2006) *Analyse de l'utilisation du métalangage linguistique par des enseignants*. Mémoire de maîtrise. Département de linguistique et de traduction, Université de Montréal.
- Picoche, J. (1993). *Didactique du vocabulaire français*. Paris : Nathan.
- Picoche, J. (2011). Lexique et vocabulaire : quelques principes d'enseignement à l'école, *Eduscol*, 1-4. [eduscol.education.fr/vocabulaire](http://eduscol.education.fr/vocabulaire)
- Polguère A. (2000). Une base de données lexicales du français et ses applications possibles en didactique, *Revue de linguistique et de didactique des langues (Lidil)*, 21, 75–97.
- Polguère, A. (2008). *Lexicologie et sémantique lexicale. Notions fondamentales*. 2e édition. Montréal : Les Presses de l'Université de Montréal, Collection « Paramètres ».
- Polguère, A. (à paraître). Terminologie grammaticale et lexicale. *Repères*.
- Polguère A. & Tremblay O. (2003) Qu'y a-t-il à l'intérieur de NOIX ? Ou comment décortiquer les unités lexicales. *La Lettre de l'AIRDF*, 33, 2003-3, 27-30.
- Psyché, V. (2007). *Rôle des ontologies dans le domaine de l'EIAH : conceptualisation, formalisation, opérationnalisation et évaluation dans le cas du design pédagogique*. Thèse de doctorat. Université du Québec à Montréal.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge : The MIT Press.
- Rastier, F. (2004). Ontologie(s). *Revue des sciences et technologies de l'information, série : Revue d'Intelligence artificielle*, 18, no 1, 15-40.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Simard, C. (1994). Pour un enseignement plus systématique du lexique. *Québec Français*, 92, 28-40.
- Simard, C. (1997). *Éléments de la didactique du français langue première*. Saint-Laurent, Québec : Éditions du renouveau pédagogique.
- Steele, J. (1990). Meaning-Text Theory and First-Language Learning : Some Implications. In J. Steele (Ed.), *Meaning-Text Theory : Linguistics, Lexicography, and Implications*. Ottawa/London/Paris, University of Ottawa Press, 377-389.
- Tremblay, Ophélie. (2009) *Une ontologie des savoirs lexicologiques pour l'élaboration d'un module de cours en didactique du lexique*. Thèse de doctorat non publiée. Département de didactique de la Faculté des sciences de l'éducation, Université de Montréal.

Tremblay, O. & Lefrançois, P. (soumis). *Pour une démarche de didactisation. Le cas des contenus théoriques en didactique du lexique*. Revue canadienne de linguistique appliquée.

Tremblay, O. & Anctil, D. (en préparation). *L'enseignement du lexique au primaire*. Montréal : Chenelière Éducation, collection « Didactique ».

---

<sup>1</sup> Pour une discussion plus détaillée sur la question (justification du choix d'une théorie linguistique de référence pour la didactique du lexique), voir Tremblay (2009, p. 34 à 58)

<sup>2</sup> Pour une liste de ces notions, voir Tremblay (2009, p. 223).

<sup>3</sup> <http://olst.ling.umontreal.ca>

<sup>4</sup> <http://protege.stanford.edu/>

<sup>5</sup> WordNet a été développé pour l'anglais, mais on en connaît des versions dans une centaine d'autres langues, répertoriées à l'adresse suivante : <http://globalwordnet.org/>.

<sup>6</sup> Un descriptif détaillé du projet se trouve à l'adresse suivante : <http://www.linguistics-ontology.org/gold.html>.

<sup>7</sup> Voir le site du projet : <http://www.lexicalmarkupframework.org/> ; la page web de la norme ISO LMF est :

[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=37327](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37327).

<sup>8</sup> Le GTN relève en fait de la classe des ontologies structurées en hiérarchies à héritage multiple (qui ne sont pas des structures véritablement arborescentes), ce qui est d'ailleurs prévu par l'environnement Protégé. Comme pour toutes les ontologies, les liens transversaux sont tissés par les *slots* (dans la terminologie Protégé) qui permettent de connecter classes et instances autrement que par le liens d'héritage. L'originalité du GTN vient quant à elle de la séparation de la classe des termes de celle des concepts linguistiques, ce qui permet notamment de gérer de façon propre une terminologie multilingue.

<sup>9</sup> Des conventions d'écriture sont nécessaires afin de distinguer concept, terme et notion. Les noms de concepts sont en anglais (il aurait aussi pu s'agir d'un numéro, puisqu'un terme précis – français et anglais dans l'état actuel de l'ontologie – est ensuite associé au concept) et ils apparaissent dans l'ontologie en majuscules.